



## Violent offenders respond to provocations with high amygdala and striatal reactivity

da Cunha-Bang, Sofi; Fisher, Patrick M.; Hjordt, Liv Vadskjær; Perfalk, Erik; Skibsted, Anine Persson; Bock, Camilla; Baandrup, Anders Ohlhues; Deen, Marie; Thomsen, Carsten; Sestoft, Dorte M.; Knudsen, Gitte M.

*Published in:*  
Social Cognitive and Affective Neuroscience

*DOI:*  
[10.1093/scan/nsx006](https://doi.org/10.1093/scan/nsx006)

*Publication date:*  
2017

*Document version*  
Publisher's PDF, also known as Version of record

*Document license:*  
[CC BY-NC](#)

*Citation for published version (APA):*  
da Cunha-Bang, S., Fisher, P. M., Hjordt, L. V., Perfalk, E., Skibsted, A. P., Bock, C., Baandrup, A. O., Deen, M., Thomsen, C., Sestoft, D. M., & Knudsen, G. M. (2017). Violent offenders respond to provocations with high amygdala and striatal reactivity. *Social Cognitive and Affective Neuroscience*, 12(5), 802-810. [nsx006].  
<https://doi.org/10.1093/scan/nsx006>

# Violent offenders respond to provocations with high amygdala and striatal reactivity

Sofi da Cunha-Bang,<sup>1,2</sup> Patrick M. Fisher,<sup>1</sup> Liv Vadskjær Hjordt,<sup>1,2</sup>  
Erik Perfalk,<sup>1,2</sup> Anine Persson Skibsted,<sup>1</sup> Camilla Bock,<sup>3</sup>  
Anders Ohlhues Baandrup,<sup>4</sup> Marie Deen,<sup>1</sup> Carsten Thomsen,<sup>2,5</sup>  
Dorte M. Sestoft,<sup>6</sup> and Gitte M. Knudsen<sup>1,2</sup>

<sup>1</sup>Neurobiology Research Unit, Copenhagen University Hospital Rigshospitalet, Blegdamsvej 9, DK-2100 Copenhagen, Denmark, <sup>2</sup>Faculty of Health and Medical Sciences, University of Copenhagen, Blegdamsvej 3B, DK-2200 Copenhagen, Denmark, <sup>3</sup>The Danish Prison and Probation Service, Herstedvester Prison, Holsbjergvej 20, DK-2620 Albertslund, Denmark, <sup>4</sup>Research Center for Advanced Imaging, Department of Radiology Zealand University Hospital, Lykkebækvej 1, DK-4600 Køge, Denmark, <sup>5</sup>Department of Radiology, Copenhagen University Hospital Rigshospitalet, Blegdamsvej 9, DK-2100 Copenhagen, Denmark, and <sup>6</sup>Ministry of Justice, Clinic of Forensic Psychiatry, Blegdamsvej 6B, DK-2200 Copenhagen, Denmark

Correspondence should be addressed to Gitte M. Knudsen, Neurobiology Research Unit and Center for Integrated Molecular Imaging, Section 6931, Rigshospitalet, Blegdamsvej 9, DK-2100 Copenhagen, Denmark. E-mail: gmk@nru.dk

## Abstract

The ability to successfully suppress impulses and angry affect is fundamental to control aggressive reactions following provocations. The aim of this study was to examine neural responses to provocations and aggression using a laboratory model of reactive aggression. We used a novel functional magnetic resonance imaging point-subtraction aggression paradigm in 44 men, of whom 18 were incarcerated violent offenders and 26 were control non-offenders. We measured brain activation following provocations (monetary subtractions), while the subjects had the possibility to behave aggressively or pursue monetary rewards. The violent offenders behaved more aggressively than controls (aggression frequency 150 vs 84,  $P = 0.03$ ) and showed significantly higher brain reactivity to provocations within the amygdala and striatum, as well as reduced amygdala-prefrontal and striato-prefrontal connectivity. Amygdala reactivity to provocations was positively correlated with task-related behavior in the violent offenders. Across groups, striatal and prefrontal reactivity to provocations was positively associated with trait anger and trait aggression. These results suggest that violent individuals display abnormally high neural sensitivity to social provocations, a sensitivity related to aggressive behavior. These findings provide novel insight into the neural pathways that are sensitive to provocations, which is critical to more effectively shaped interventions that aim to reduce pathological aggressive behavior.

**Key words:** aggression; connectivity; fMRI; point subtraction aggression paradigm; PSAP; psychopathy

Received: 17 August 2016; Revised: 11 November 2016; Accepted: 16 January 2017

© The Author (2017). Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com)

## Introduction

Aggression and impulsivity play a critical role in the manifestation of violent and criminal behaviors, thereby posing large costs to the victims and the society. Delineating relevant neural pathways represents a critical step towards developing more effective preventative approaches and therapeutic strategies for curtailing pathological aggressive behavior.

Existing neuroimaging studies in patient groups with high levels of aggression and in healthy controls implicate a network of brain regions involved in aggression that includes the amygdala (McCloskey et al., 2016), striatum (Glenn and Yang, 2012), anterior cingulate and orbitofrontal cortex (Beyer et al., 2014). The dominant conceptual framework of how these regions regulate reactive aggression is that the prefrontal cortex inhibits or modulates subcortical activity mediating the aggressive response (Nelson and Trainor, 2007; Siever, 2008; Rosell and Siever, 2015). That is, reduced prefrontal activity combined with heightened subcortical activity in the context of threat-related or provocative stimuli poses an increased risk for impulsive aggression. Such theory is supported by reduced amygdala-orbitofrontal coupling in patients with intermittent explosive disorder relative to controls when presented to angry faces (Coccaro et al., 2007), and reduced resting-state amygdala-ventromedial prefrontal connectivity in psychopathic inmates compared to inmates with low psychopathic traits (Motzkin et al., 2011).

However, few functional neuroimaging studies have used tasks directly targeting reactive aggression in samples of pathologically aggressive individuals. Typically, brain responses to aversive facial expressions are used. For example, the involvement of the amygdala in aggression has been evaluated in patients with intermittent explosive disorder, who display a heightened amygdala response to angry faces (Coccaro et al., 2007; McCloskey et al., 2016), and in youths with conduct disorder and callous-unemotional traits, who show reduced amygdala reactivity to fearful faces (White et al., 2012).

We recently implemented a laboratory model of reactive aggression for use in fMRI; the point subtraction aggression paradigm (PSAP). The PSAP is a paradigm wherein a fictitious opponent periodically steals money with the aim to provoke the participant, who can choose to act aggressively or pursue monetary rewards (Cherek et al., 1997). In healthy controls, provocations (monetary subtractions from the fictitious opponent) activated several key brain regions implicated in aggressive behavior including the amygdala, anterior cingulate cortex, insula, striatum and orbitofrontal cortex (Skibsted et al., 2015). The PSAP in fMRI thus constitutes a useful instrument for investigating neural pathways underlying aggression in humans but has until now never been employed in individuals with extreme levels of aggression or a documented history of impulsive violent behavior.

In this study, we examine neural responses to provocations and aggressive behavior in violent offenders and in healthy control non-offenders using the PSAP during fMRI. Our first hypothesis was that during provocations and aggressive behavior, violent offenders would show higher blood-oxygen level dependent (BOLD) activity in subcortical regions (the amygdala and striatum), and reduced functional connectivity between the amygdala and the prefrontal cortex. This hypothesis was based on aggression being a categorical construct (violent offender or not). Alternatively, aggression can also be viewed as an individually varying personality trait (i.e. level of aggression). Therefore, our second hypothesis was that BOLD activity during

provocations and/or aggressive behavior would be positively associated with how often participants were aggressive during the PSAP task and with self-reported trait anger and aggression.

## Materials and methods

### Participants

The final study sample consisted of 18 incarcerated violent offenders with a documented history of severe violent crimes and 26 healthy control subjects recruited from the community. The participants were all male and groups were age-matched. Violent offenders were recruited from closed state prisons within the National Prison and Probation Service in Denmark. Inmates who had a documented history of convictions for violent crimes (of which at least one was impulsive in nature, e.g. homicide, rape, aggravated assault, uttering threats) were invited to an initial screening interview conducted by a medical doctor and a psychologist. Healthy control non-offenders were recruited via community websites and bulletin boards in vocational schools. Only men were included consistent with the high predominance of male inmates in the prisons.

Exclusion criteria for all participants were current or lifetime history of major psychiatric disorders (major depressive disorder, bipolar disorder or psychotic symptomatology), symptomatic medical or neurological illness, severe head trauma, severe visual or hearing impairment, contraindications for MRI, use of psychotropic medications, current substance or alcohol abuse. Fourteen violent offenders had a history of substance abuse including cannabis ( $n = 10$ ), cocaine ( $n = 7$ ), alcohol ( $n = 5$ ), stimulants ( $n = 2$ ), opioids ( $n = 3$ ) and anabolic steroids ( $n = 2$ ), but all had been in remission for at least six months. All participants tested negative on urine drug screen (Rapid Response Multi-Drug; BTNX Inc., Toronto, Ontario, Canada) on the day of scanning and had an unremarkable MRI. As evaluated on the day of scanning, none of the participants had any significant medical or neurological illness according to the Schedules for Clinical Assessment in Neuropsychiatry version 2.1, physical examination or blood biochemistry.

Fifty-one subjects were initially enrolled for the study, after which seven subjects were excluded: one (healthy control) did not believe the deception of the paradigm, one (healthy control) with a pathological MRI, two (one healthy control and one violent offender) who became claustrophobic and two violent offenders who reported taking psychotropic medications at the time of scanning. One healthy control was excluded due to aberrant behavior, indicating a misunderstanding of the instructions. Forty-three participants also participated in a positron emission tomography study reported elsewhere (da Cunha-Bang et al., 2016).

The study was approved by the National Prison and Probation Service and the local ethical committee (Copenhagen, Denmark, reference H-3-2013-100). All participants provided written informed consent following full description of the procedures, which for the violent offenders included access to criminal files, and received monetary compensation for their participation.

### Personality assessment

The Structured Clinical Interview for DSM-IV II (SCID-II) was administered for assessment of personality disorders. All violent offenders were diagnosed with one or multiple personality disorders; antisocial ( $n = 14$ ), borderline ( $n = 2$ ), schizoid ( $n = 1$ ), dependent ( $n = 1$ ),

paranoid ( $n=1$ ), obsessive-compulsive ( $n=1$ ) and unspecified personality disorder ( $n=7$ ). Two medical doctors administered the SCID-II, and co-rating was conducted on a subset of the violent offenders, which yielded full final diagnostic consistency. Level of psychopathy was assessed using the Psychopathy Checklist-Revised (PCL-R) (Hare, 2003), which consists of 20 items scored from 0 to 2 based on the presence of each trait. PCL-R interviews were conducted by course-certified health professionals with a medical or psychology background, and all ratings were consensus decisions based on notes from each interview and collateral information from criminal files. Intelligence quotient (IQ) was evaluated by a trained neuropsychologist using the Reynolds Intellectual Screening Test (RIST). Four healthy control participants did not undergo the RIST.

To assess trait aggression, trait anger and trait impulsivity, we used the Buss-Perry Aggression Questionnaire (BPAQ) (Buss and Perry, 1992; da Cunha-Bang et al., 2013), the State-Trait Anger Expression Inventory 2 (STAXI-2) (Moeller et al., 2015) and the Barratt's Impulsiveness Scale version 11 (BIS) (Patton et al., 1995; da Cunha-Bang et al., 2013), respectively. One healthy control did not complete the STAXI-2.

### fMRI paradigm

Prior to scanning, participants were instructed that they would play a game with another study participant and they could earn points, which could be exchanged for money. Participants completed one 12-minute session of the PSAP and the monetary reward was 1.34 Euro (10 DKK) per point won. The participants responded via a five-finger button response unit on the right hand. Three response options were available: pressing the button for 'Option 1' 100 times resulted in the participant earning 1 point (1.34 Euro), pressing the button for 'Option 2' 10 times resulted in removal of a point from the opponent and pressing the button for 'Option 3' 10 times briefly protected the participant's point total from the opponent stealing. Options 1, 2 and 3 corresponded to the index, middle and ring finger keys, respectively. Game status was continuously presented on a screen viewable by participants while in the scanner (Figure 1A). This included the number of points currently earned and the number of button presses currently executed within the current chosen option, which was highlighted red. Participants were made aware of earning a point when the button press counter reached 100, which elicited black flashing positive symbols ("+" around the point total (lasting for 1000 ms with "+" symbols shown three times for 200 ms each time, Figure 1D) that then increased by one point. When points were stolen from the participant, this was indicated by a similarly flashing red negative symbols ("−") around the point total, which then decreased by one point (lasting for 1000 ms with "−" symbols showing three times for 200 ms each time, Figure 1C). Participants had to complete a started option before choosing a new option (unselected options were grayed until the chosen option was completed). Participants were provoked by having points stolen from them pseudo-randomly within a time frame of 6–60 s in the absence of using Option 2 or 3. After completing Option 2 or 3, a provocation-free interval of pseudo-random length (max 60 s) was initiated. However, participants were only aware of the potential protective effect of Option 3.

Participants were informed that they were assigned to the experimental condition in which they did not keep the points they stole from the opponent. However, they were also told that the opponent was allowed to keep the points he stole from the participant. In this way, the participant choosing Option 2 reflects an aggressive behavior, void of any monetary incentive, the opponent choosing Option 2 represents a provocative stimulus.

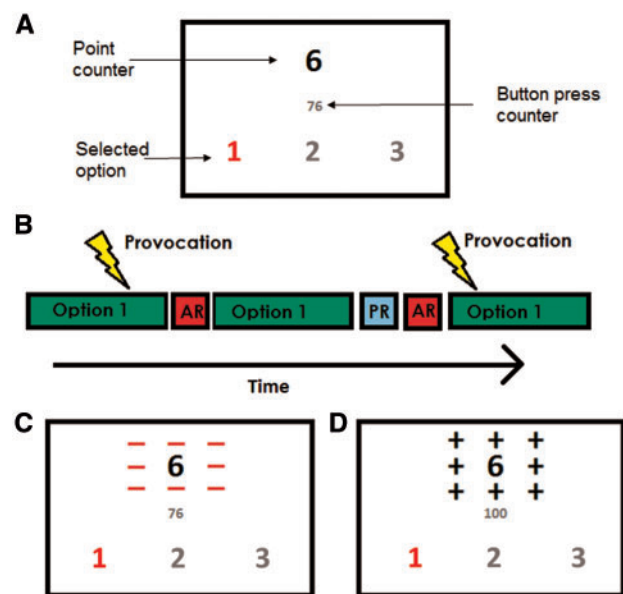


Fig. 1 (A) Screen displaying what the participants viewed in the scanner. The red-colored digit denotes that the participant is currently in Option 1. (B) Timeline with schematic representation of the task conditions. AR, aggressive response; PR, protective response. (C) Screen displaying a provocation (the opponent steals a point, negative symbols flash around the point counter). (D) Screen displaying the participants winning a point (positive symbols flash around the point counter).

Participants completed a 1-min trial session in the scanner immediately before playing one 12-min session of the PSAP. Participants completed a questionnaire immediately after the scan session and outside the scanner to determine their impression of the opponent. Participants who indicated they thought they played against a computer were excluded.

### Imaging acquisition

MRI scans were acquired on a 3T Prisma scanner (Siemens, Erlangen, Germany) with a 64-channel head coil. For blood oxygen level dependent (BOLD) fMRI, a T2\*-weighted gradient echo-planar imaging (EPI) sequence was used with a repetition time of 2000 ms, echo time of 30 ms, flip angle of 90°, and 32 slices with a slice thickness of 3.0 mm (0.75 mm gap). A total of 360 whole-brain fMRI volumes were acquired. We acquired a T1-weighted, TurboFLASH sequence, high-resolution whole-brain three-dimensional structural magnetic resonance scan with an inversion time of 900 ms, echo time of 2.58 ms, repetition time of 1900 ms, flip angle of 9°, in-plane matrix of 256 × 256, in-plane resolution of 0.9 × 0.9 mm, 224 slices and a slice thickness of 0.9 mm, no gap. To reduce motion, an in-house made head fixation system was used.

### fMRI data analysis

Functional neuroimaging data were analyzed with SPM8 (<http://www.fil.ion.ucl.ac.uk/spm>). Single-subject functional images were spatially realigned to the first image. The T1-weighted structural image was co-registered to the first functional image and the origin reset to the anterior commissure (AC) using *acpcdetect* (<https://www.nitrc.org/projects/art>). The co-registered T1-weighted image was normalized into Montreal Neurological Institute (MNI) stereotactic space and the normalization parameters were applied to the functional images. Normalized functional images were smoothed using an 8-mm



FWHM Gaussian filter. We used Artifact Detection Tools ([http://www.nitrc.org/projects/artifact\\_detect](http://www.nitrc.org/projects/artifact_detect)) to identify individual functional volumes that deviated significantly from the subject specific dataset in terms of motion or signal variability. Flagged volumes were censored when estimating task-related effects. Only one participant moved greater than three millimeters in any direction and only at two distinct time points. These volumes were censored with Artifact Detection Tools.

We defined the following conditions (Figure 1): Option1 (block, the first ten seconds of Option1), Aggressive Response (block, duration of Option2) and Provocation (event, at time of provocation). Only the first 10s of Option 1 was used as the “baseline condition” to limit potential reward-related brain responses. If a provocation occurred during the first 10s of Option 1, the time from beginning of Option 1 until the provocation occurred was used. We used this as a baseline condition because it is more neutral than for example using an event when the participants win a point (that would be visually more equivalent to provocations). The total duration of “baseline condition” did not significantly differ between violent offenders (mean: 211.2, standard deviation: 35.6) and healthy controls (mean: 226.4 standard deviation: 26.0),  $P=0.1$  (two-sample t-test). The following contrasts of interest were then estimated: Provocations > Option 1 and Aggressive Response > Option 1.

Single-subject design matrices were estimated using the general linear model to determine condition-specific BOLD responses. Individual contrast images (i.e. weighted sum of beta images) were included in group-level analyses to determine task-related brain responses using one-sample t-tests. Given the strong *a priori* evidence, analyses focused on the following regions of interest (ROIs) defined within WFU Pickatlas: the amygdala, striatum (i.e. caudate and putamen) and the prefrontal cortex (including the anterior cingulate cortex). We used the SPM anatomy toolbox version 1.5 to create anatomical probability maps for amygdala sub-regions; the superficial, centromedial and laterobasal complex (Amunts et al., 2005). To address the issue of multiple comparisons we used 3dClustSim (version 8 July 2016, which accounts for a previously reported threshold bug in 3dClustSim, Eklund et al., 2016), a software program within AFNI (<http://afni.nimh.nih.gov/afni>) that uses a Monte Carlo simulation method to establish family-wise-error-corrected cluster extent thresholds unlikely to have occurred by chance ( $\alpha < 0.05$ ; Forman et al., 1995). We used a voxel-level statistical threshold of  $P < 0.001$ , uncorrected. A recent study indicates that this threshold reasonably controls for the family-wise error rate (Eklund et al., 2016). Based on this, amygdala, striatum, prefrontal and wholebrain clusters of  $k \geq 1$ , 18, 113 and 175 voxels, respectively, were considered statistically significant. Coordinates are reported in Montreal Neurological Institute space as  $[x,y,z]$  and  $k$  denotes cluster sizes.

We extracted mean signal values for significantly activated clusters within each ROI bilaterally across all participants and used these values for statistical analyses using R version 3.1.1 (<https://cran.r-project.org/>). Group differences in demographic data, task behavior and brain responses were determined using two-sample t-tests. When data were non-normally distributed, we also evaluated group differences using non-parametric tests (Wilcoxon). In all instances, the results were similar and we, therefore, present results using only t-tests for clarity. To investigate whether brain responses were reflected in the behavioral data, linear regression analyses were employed to evaluate the association between task-related behavior, trait anger, trait aggression and trait impulsivity and neural responses to provocations, including age and group as covariates. We also evaluated whether group moderated the associations between brain responses and behavioral data (interaction

analyses). To account for differences in number of provocations and number of total button presses, we included these covariates in the analyses with task-related behaviors. In the violent offenders, we also evaluated associations between psychopathy score (PCL-R) and reactivity to provocations.

In a post-hoc sensitivity analysis of the observed group differences in amygdala and striatal reactivity to provocations, we evaluated group differences in multiple linear regression analyses, including IQ, number of provocations, total button presses and duration of baseline condition as covariates (Supplementary Tables 1 and 2). To ensure that extracting values from significantly activated clusters across two groups did not bias the results, we also evaluated the findings (reactivity to provocations) using extracted mean signal values from anatomically defined ROIs across participants (Supplementary Tables 3–6).

## Psychophysiological interaction analyses

We used psychophysiological interaction (PPI) analyses (Friston et al., 1997; McLaren et al., 2012; Madsen et al., 2015) to assess the provocation-associated functional connectivity with seed regions in either amygdala or striatum, using the contrast Provocations > Option 1. Bilateral amygdala and striatum seeds were defined based on clusters significantly responsive to provocations across subjects (Table 2). Although our *a priori* hypothesis included only amygdala connectivity, we probed bilateral striatum connectivity following our observed association between trait anger and striatal reactivity to provocations. The mean seed timeseries for each subject was extracted using the generalized PPI toolbox v7.12 (McLaren et al., 2012). Single-subject design matrices estimated to determine PPI effects were identical to design matrices used to estimate main effects of task except including the seed timeseries and psychophysiological interaction terms as additional regressors. Individual contrast images were included in group-level analyses for evaluation of whole-brain connectivity with the bilateral amygdala or striatum as seeds. Seeds were analyzed independently. A two-sample t-test within SPM was used to identify group differences in connectivity using the statistical thresholds described above. Post hoc, we confirmed our connectivity findings using anatomically defined amygdala and striatal seeds, presented in Supplementary Figure 2.

## Results

### Participant characteristics and task behavior

Demographic, clinical and behavioral characteristics of the participants are provided in Table 1. Sixteen violent offenders met European criteria (Cooke et al., 2005) for psychopathy (PCL-R > 25), whereas only nine of those individuals met an alternative threshold (PCL-R  $\geq 30$ ) for psychopathy (Hare, 2003). The mean ( $\pm$  standard deviation) PCL-R total score was  $29.3 \pm 4.3$  (mean factor 1 score  $12.8 \pm 1.8$  and mean factor 2 score  $14.1 \pm 3.6$ ). Behaviorally, violent offenders used the aggressive response in the PSAP twice more frequently than controls (mean Option 2 presses  $150.0$  vs  $84.0$ ,  $P = 0.03$ ). Eight control subjects and only one violent offender did not use the aggressive response. Thus, results of the aggressive response are based on 35 participants (17 violent offenders and 17 healthy controls).

### Functional imaging data

**Main task effects.** ROI analyses across all participants revealed significant activation within the amygdala, striatum and prefrontal

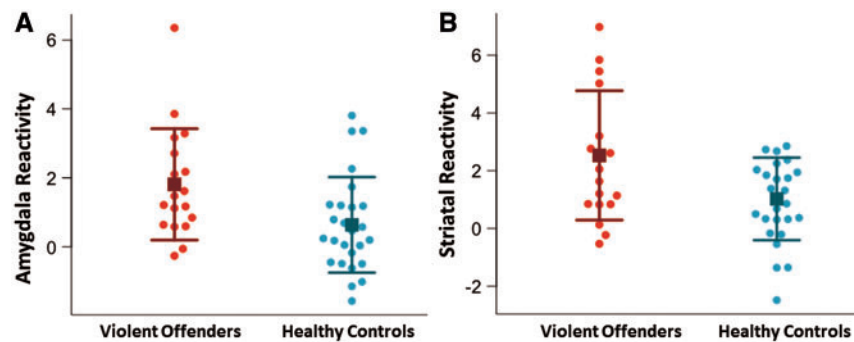


Fig. 2. Heightened amygdala and striatal reactivity to provocations in violent offenders. (A) "Amygdala reactivity" represents signal values extracted from left and right amygdala clusters significantly activated in response to provocations across all participants. Violent offenders show significantly higher amygdala reactivity to provocations ( $P=0.02$ ). (B) "Striatal Reactivity" represents signal values extracted from left and right striatal clusters significantly activated in response to provocations across all participants. Violent offenders show significantly higher amygdala reactivity to provocations ( $P=0.02$ ). Squares represent group mean and error bars represent standard deviations.

Table 1. Demographics, personality and task behavior

	Violent offenders	Healthy controls	P value
Number of subjects	18	26	
Age, years	31.8 ± 8.8	29.6 ± 9.2	0.4
Duration of education, years	8.9 ± 2.6	11.6 ± 0.8	0.0004
IQ, RIST score	99.1 ± 8.1	108.9 ± 6.9	0.0003
Number of violent convictions	3.7 ± 2.4	None	
Number of violent charges against	16.4 ± 30.2	None	
Age at first violent conviction, years	19.5 ± 4.3	n/a	
Personality traits			
Trait aggression <sup>a</sup>	87.1 ± 22.2	54.3 ± 15.9	0.00009
Trait anger <sup>b</sup>	21.0 ± 6.8	14.6 ± 2.3	0.001
Trait impulsivity <sup>c</sup>	66.7 ± 10.0	60.9 ± 9.4	0.07
PSAP behavior			
Option 1	2283.6 ± 378.9	2476.0 ± 276.6	0.08
Option 2	150.0 ± 98.0	84.0 ± 95.2	0.03
Option 3	175.0 ± 11.5	258.4 ± 94.4	0.01
Total button presses	2608.6 ± 304.8	2818.4 ± 170.7	0.01
Provocations	10.6 ± 1.6	10.2 ± 2.0	0.5
Points earned	12.1 ± 3.4	14.3 ± 3.2	0.04
Option 2 per provocation	14.5 ± 9.5	8.7 ± 9.9	0.06

RIST, Reynolds Intellectual Screening Test; IQ, intelligence quotient; PCL-R, psychopathy checklist revised. P values represent two-sample t-tests.

<sup>a</sup>Buss-Perry Aggression Questionnaire total score.

<sup>b</sup>State-Trait Anger Expression Inventory, Trait Anger Scale total score ( $n=40$ ).

<sup>c</sup>Barratt's Impulsiveness Scale Version 11 total score.

cortex during provocations (Table 2). Of the 46 functionally activated voxels in the right amygdala, 37 voxels were located within the superficial complex and nine voxels within the laterobasal complex. Of the 10 functionally activated voxels in the left amygdala, five were located within the laterobasal complex and five within the superficial nuclei. We found no significant main effects of task during the aggressive response in any ROIs.

### Group comparisons

ROI analyses revealed that violent offenders had significantly higher reactivity to provocations bilaterally in amygdala (mean

Table 2. Task-related reactivity to provocations (provocations > Option 1) within regions of interest across all participants. Clusters reflect a voxel-level significance threshold of  $P < 0.001$ , uncorrected, with region specific cluster extent thresholds. Coordinates reported in Montreal Neurological Institute space

Anatomical region	Cluster size	MNI coordinates of peak voxel			Z-score
		X	Y	Z	
Amygdala	46	20	-4	-16	4.86
	10	-22	0	-18	3.31
Striatum	213	-8	8	2	5.88
	232	10	8	2	5.72
Prefrontal cortex	4356	32	24	-6	7.71
	1853	4	30	44	6.59
	1005	-42	10	26	6.51
	416	-30	24	-10	6.45
	201	-32	-4	52	5.05

signal value in violent offenders: 1.8, healthy controls: 0.6,  $t=2.5$ ,  $df=32.9$ ,  $P=0.02$ , Figure 2) and bilaterally in striatum (mean signal value in violent offenders: 2.5, healthy controls: 1.0,  $t=2.5$ ,  $df=26.5$ ,  $P=0.02$ , Figure 2). The groups did not differ significantly in response to provocations within the prefrontal cortex ( $P=0.5$ ). The observed group differences remained significant after including IQ, number of provocations, total button presses and duration of baseline condition as covariates (Supplementary Table 1), as well as using anatomical definitions of the ROIs (Supplementary Table 3). The group difference in amygdala reactivity to provocations also remained statistically significant after exclusion of a violent offender with very high activity (mean signal value in violent offenders: 1.5, healthy controls: 0.6,  $t=2.3$ ,  $P=0.03$ , Supplementary Figure 1).

### Brain responses to provocations and task-related behavior

The association between amygdala reactivity to provocations and task-related aggressive behavior was significantly moderated by the group (difference in slopes: -40.7, standard error (SE): 19.3, 95% confidence interval (CI): [-79.8; -1.6],  $P=0.04$ , effect in violent offenders: 37.2 aggressive button presses per unit

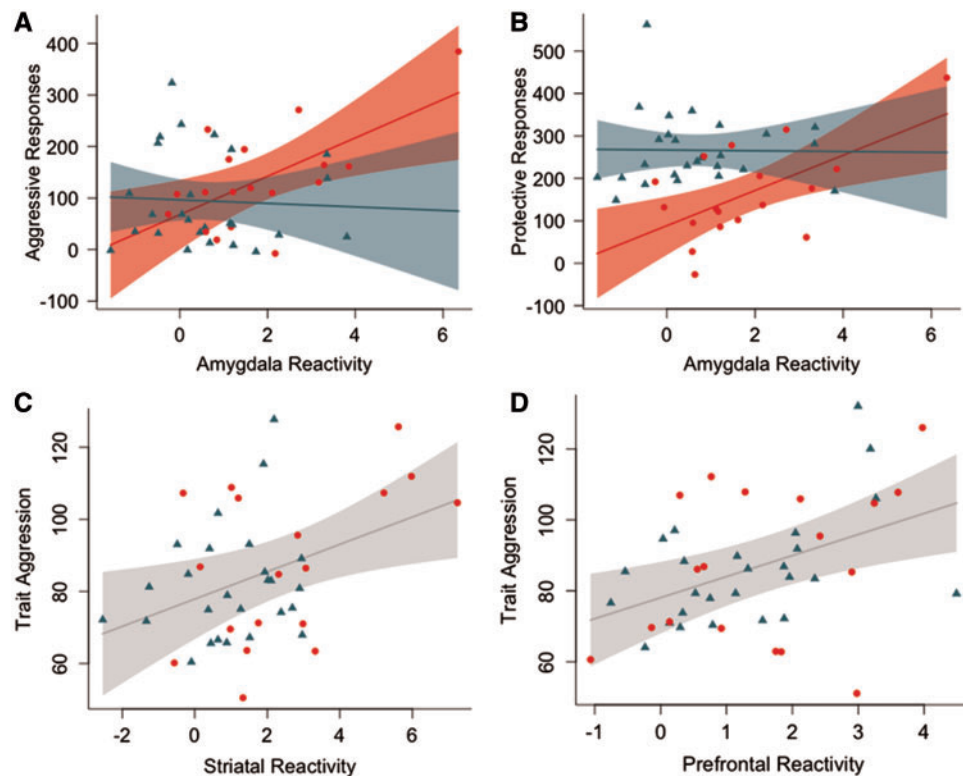


Fig. 3. The associations between amygdala reactivity to provocations and task-related aggressive behavior (A) and protective behavior (B) were significantly moderated by group. Plots (A) and (B) are shown given a mean age, number of button presses and number of provocations. Associations between striatal (C) and prefrontal (D) reactivity to provocations and trait aggression (Buss-Perry Aggression Questionnaire total score) across all participants. Plots (C) and (D) are shown given a mean age. Red circles represent violent offenders; blue triangles represent healthy controls. Shaded areas represent pointwise 95% confidence interval of fit line.

change in amygdala reactivity, SE: 13.9, CI: [9.1;65.4],  $P=0.01$ , effect in healthy controls:  $-3.4$  aggressive button presses per unit change in amygdala reactivity, SE: 13.3, CI:  $[-29.5;22.6]$ ,  $P=0.8$ , Figure 3). Task-related aggressive behavior was not associated with striatal or prefrontal reactivity to provocations.

We also found that group significantly moderated the association between the number of protective button presses and amygdala reactivity to provocations (difference in slopes:  $-42.4$ , SE: 41.5, CI:  $[-82.0;2.8]$ ,  $P=0.04$ , effect in violent offenders:  $41.5$  protective button presses per unit change in amygdala reactivity, SE: 14.1, CI:  $[13.0;70.0]$ ,  $P=0.005$ , effect in healthy controls:  $-0.9$  protective button presses per unit change in amygdala reactivity, SE: 13.4, CI:  $[-27.3;25.5]$ ,  $P=0.9$ , Figure 3), and striatal reactivity to provocations (difference in slopes:  $-45.1$ , SE: 16.4, CI:  $[-78.3; -12.0]$ ,  $P=0.009$ , effect in violent offenders:  $28.5$ , SE: 10.3, CI:  $[7.6;49.3]$ ,  $P=0.009$ , effect in healthy controls:  $-16.7$ , SE: 13.0, CI:  $[-42.3;8.8]$ ,  $P=0.2$ ) and prefrontal reactivity to provocations (difference in slopes:  $-52.2$ , standard error: 22.2, CI:  $[-97.2; -7.3]$ ,  $P=0.02$ , effect in violent offenders:  $34.7$ , SE: 16.5, CI:  $[1.4;68.1]$ ,  $P=0.04$ , effect in healthy controls:  $-17.5$ , SE: 15.4, CI:  $[-47.6;12.6]$ ,  $P=0.25$ ). Associations between reactivity to provocations within anatomically defined ROIs and task-related behavior are presented in Supplementary Tables 4 and 5.

### Brain responses to provocations and aggressive traits

Across all participants, striatal reactivity to provocations was positively correlated with trait aggression (slope estimate: 3.8, standard error: 1.5, 95% CI:  $[0.7;6.8]$ ,  $P=0.02$ , Figure 3) and with trait anger (slope estimate: 1.2, standard error: 0.4, 95% CI:  $[0.4;2.0]$ ,  $P=0.005$ ). Prefrontal reactivity to provocations was positively

correlated with trait aggression (slope estimate: 6.0, standard error: 2.0, 95% CI:  $[1.9;10.0]$ ,  $P=0.005$ , Figure 3), and with trait anger (slope estimate: 1.2, standard error: 0.6, 95% CI:  $[0.05;2.3]$ ,  $P=0.04$ ). The effect of covariates (IQ, number of provocations, button presses and duration of baseline condition) on these associations is presented in Supplementary Table 2, and associations between reactivity within anatomically defined ROIs and aggressive traits are presented in Supplementary Table 6. BIS or PCL-R scores were not associated with amygdala, striatal or prefrontal reactivity to provocations. Group did not moderate the association between brain responses to provocations and aggressive traits.

### Functional connectivity

The PPI analysis of functional connectivity with seeds placed in the amygdala revealed that when provoked, violent offenders had significantly reduced connectivity between the amygdala and a cluster covering the right superior prefrontal gyrus ( $k=527$  voxels,  $[18,42,32]$ ,  $z=3.78$ , Figure 4). Violent offenders also had significantly reduced connectivity between the striatum and a bilateral prefrontal cluster covering the medial orbitofrontal cortex and medial anterior cingulate cortex ( $k=349$  voxels,  $[-10,40, -8]$ ,  $z=3.69$ , Figure 4). In the violent offenders, we also observed reduced functional connectivity between prefrontal regions and anatomically defined amygdala and striatum seeds (Supplementary Figure 4).

### Discussion

This study is the first to investigate brain reactivity to provocations in aggressive violent offenders using the PSAP. The most



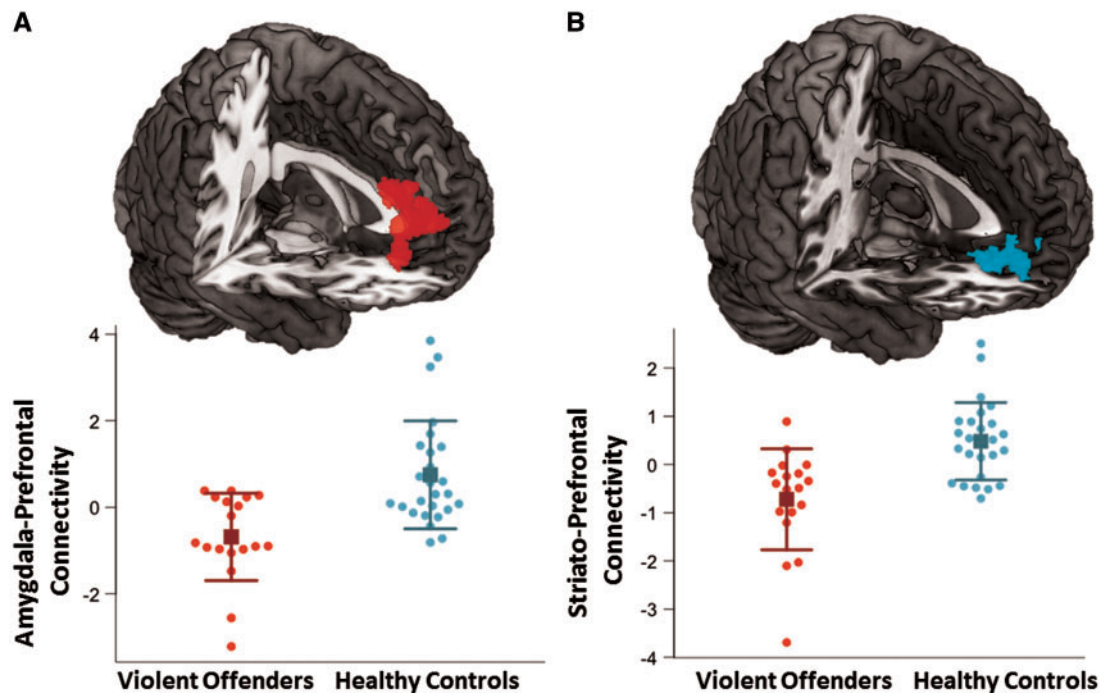


Fig. 4. Group differences in amygdala and striatal functional connectivity in response to provocations. (A) Cluster in which bilateral amygdala functional connectivity was significantly greater in control subjects relative to violent offenders ( $k = 527$  voxels,  $[18, 42, 32]$ ,  $z = 3.78$ ). The corresponding plot below represents the extracted mean signal values from this cluster with means (squares) and standard deviations (error bars). (B) Cluster in which the bilateral striatum functional connectivity was significantly greater in control subjects relative to violent offenders ( $k = 349$  voxels,  $[-10, 40, -8]$ ,  $z = 3.69$ ). The corresponding plot below represents the extracted mean signal values from this cluster with means (squares) and standard deviations (error bars).

striking findings are the heightened neural response to provocations in violent offenders within the amygdala and striatum, and the reduced functional connectivity between these regions and the prefrontal cortex in the context of provocations. Our data support a model where reactive aggression following provocation emerges as the result of heightened subcortical reactivity and reduced fronto-limbic connectivity. This interpretation is further substantiated by the neural activation patterns during provocations; these correlated positively with task-related behavior and self-report measures of trait aggression and trait anger. Taken together, these findings highlight critical neural pathways underlying variability in aggression-related behavioral phenotypes.

Given that the violent offenders in the current study had intermediate to high levels of psychopathy, the heightened amygdala reactivity to provocations may seem contradictory with previous reports of amygdala hypo-responsivity to indices of threats in individuals with psychopathy (Blair, 2010). However, such hypo-responsivity to threat-related stimuli has primarily been linked to cues of fear or distress whereas people scoring high in psychopathic traits (lifestyle facet) show amygdala hyper-reactivity to interpersonal signals of threat such as angry faces (Carre et al., 2013). Also, reactive aggression in psychopathy is proposed to be frustration-based, as opposed to threat-related (Blair, 2010, 2012). Frustrations can occur when another person's behavior undermines achieving an expected reward/goal (Blair, 2012). Having points/money stolen while working for a monetary reward in the PSAP putatively evoke similar frustrations and anger. Indeed, when rewards were blocked in an fMRI monetary reward paradigm, amygdala reactivity correlated positively with frustration (Yu et al., 2014), and in an economic exchange paradigm, amygdala reactivity to unfair offers correlated positively with offer

rejection (Gospic et al., 2011). The present study provides insight into the functional role of the amygdala in pathological aggression, particularly in terms of its putative differential function in reactive and instrumental aggression. It has been proposed that amygdala responding can serve to distinguish instrumental forms of aggression associated with antisocial personality disorder and psychopathy from reactive aggression, which is more characteristic of borderline personality disorder and intermittent explosive disorder (Coccaro et al., 2011). However, individuals with psychopathy are at increased risk for both instrumental and reactive aggression (Blair, 2007, 2010). Amygdala hyper-responsiveness to negative emotional stimuli and social threats has been observed in several studies of patients with borderline personality disorder (Schulze et al., 2015) and intermittent explosive disorder (Coccaro et al., 2007; McCloskey et al., 2016) but reactive aggression in psychopathy has received less attention. Here we find evidence for a similar positive association between amygdala reactivity to provocations and aggression in individuals with intermediate-to-high levels of psychopathy. In the present study, violent offenders were selected based on a history of impulsive violent crimes, but most violent offenders had also committed numerous crimes more instrumental in character (e.g. robbery and fraud), emphasizing that these two types of aggression often coexist. Although not possible within the current study design, future studies employing tasks assaying both forms of aggression within the same cohort would shed light on the differential role of relevant neural circuits in reactive and instrumental aggression and psychopathy.

In addition to the amygdala, violent offenders also responded to provocations with higher activity within the dorsal striatum. The striatum is implicated in a variety of functions, including motor function, motivated behaviors and reward



(Zink et al., 2004). However, several types of non-rewarding stimuli activate the striatum and it is suggested that the striatum encodes all salient stimuli (Zink et al., 2003). Striatal reactivity to provocations might reflect the saliency of this stimulus, but we also speculate that striatal reactivity to provocations might reflect motor vigilance. Importantly, the observation that striatal reactivity to provocations correlated positively with trait aggression and trait anger substantiates the involvement of the striatum in aggression. Recent evidence implicates a role for the striatum in antisocial behavior and psychopathy (Glenn and Yang, 2012; Yang et al., 2015). Increased striatal volume is reported in adults with psychopathy (Glenn et al., 2010), antisocial personality disorder (Barkataki et al., 2006) and in adolescents with psychopathic traits (Yang et al., 2015). In human functional neuroimaging studies, the impulsive-antisocial factor scores of self-reported psychopathic traits in healthy subjects correlated positively with amphetamine-induced dopamine release using positron emission tomography and with the BOLD response to reward anticipation in the nucleus accumbens (Buckholz et al., 2010). A dysfunction of the striatum is speculated to contribute to the increased sensation-seeking, reward-driven and impulsive behavior in antisocial personality disorder and psychopathy (Yang et al., 2015). Our finding that violent offenders show heightened striatal reactivity to provocations supports an aberrant striatal function in individuals with high levels of aggression in response to aversive salient stimuli.

One theory holds that amygdala activity is controlled by prefrontal inhibitory projections and that this process is disrupted in aggressive individuals (Rosell and Siever, 2015). In resting-state fMRI, it was shown that 20 inmates fulfilling criteria for psychopathy had reduced amygdala-ventromedial prefrontal cortex connectivity compared with non-psychopathic inmates (Motzkin et al., 2011), and an fMRI study using the ultimatum game in 30 youths with disruptive behavior disorders reported reduced amygdala-prefrontal connectivity during high provocation trials compared with controls (White et al., 2015). Our findings support that violent offenders have reduced amygdala-prefrontal connectivity in the context of provocations, possibly reflecting reduced top-down prefrontal regulation. These findings strongly corroborate the observation in disruptive behavior-disordered youths and are consistent with the model of impaired prefrontal regulation of emotional reactions to provocations in antisocial individuals. Collectively, our data point toward a distributed neural circuit wherein the combination of heightened amygdala and striatal reactivity to provocations (possibly stemming from reduced prefrontal regulation) shapes heightened aggressive behaviors.

An important strength of this study was that the activation patterns in response to provocations could be related to behavior within the paradigm as well as to trait aggression and anger. However, the study presents with limitations inherent to studying criminal violent offenders, such as group differences in IQ and education, previous drug abuse, smoking and incarceration. We performed supplementary analyses with the inclusion of IQ as a covariate, which did not change the main outcomes (Supplementary material). With respect to diagnoses, the violent offenders presented with mixed personality disorders, so the results cannot be attributed to specific personality pathology but rather to the participants' aggressive and violent behavior. Further, we confirmed our results based on functionally activated clusters across groups in an independent analysis based on anatomically defined clusters. Lastly, it should be noted that the PSAP is a paradigm that is user driven, i.e. the

number of provocations are not fixed and the participants are not forced to use the aggressive response. Although this means that participants can choose not to behave aggressively, it also makes the paradigm more translatable to a realistic setting. Importantly, including the number of provocations (ranging from 8 to 16) and total button presses as covariates did not change the main outcome (Supplementary Tables 1 and 2).

In conclusion, we find that violent offenders display abnormally high neural reactivity to provocations within the amygdala and striatum, and that this sensitivity is related to aggressive behavior. We also demonstrate that violent offenders show reduced amygdala-prefrontal and striato-prefrontal connectivity in the context of provocations. These data provide novel evidence of aberrant brain function in a unique cohort of individuals with a history of extremely violent behavior. The findings suggest that an exaggerated neurobiological sensitivity to provocations or frustrations and lack of prefrontal control are key features of pathological aggression. Prevention and treatment of aggressive behaviors would benefit from interventions targeting this type of vulnerability.

## Supplementary data

Supplementary data are available at SCAN online.

## Acknowledgements

We thank all the volunteers for kindly participating in this study. The excellent technical assistance of Lone Ibsgaard Freyr, Martin Korsbak Madsen and Gerda Thomsen is gratefully acknowledged. We also wish to thank all prison staff for the excellent collaboration and Vibeke Dam for assistance with neuropsychological testing.

## Funding

This work was supported by the Danish Council for Independent Research [Grant no. 1331-00328] and Rigshospitalets Research Council [Grant nos. R49-A1646 and R65-A2250]. The funding sources were not involved in the study design or in the collection, analysis, writing or publication of data.

## Conflicts of interests

G.M.K. has received honoraria as a consultant for H. Lundbeck A/S, as a member of the steering group for Brain Prize. She is also on the advisory board for the Kristian G. Jebsen Foundation and a field editor for International Journal of Neuropsychopharm. All other authors declare no conflicts of interest and report no financial disclosures.

## References

- Amunts, K., Kedo, O., Kindler, M., et al. (2005). Cytoarchitectonic mapping of the human amygdala, hippocampal region and entorhinal cortex: intersubject variability and probability maps. *Anatomy and Embryology (Berlin)*, 210(5–6), 343–52.
- Barkataki, I., Kumari, V., Das, M., Taylor, P., Sharma, T. (2006). Volumetric structural brain abnormalities in men with schizophrenia or antisocial personality disorder. *Behavioural Brain Research*, 169(2), 239–47.

- Beyer, F., Munte, T.F., Gottlich, M., Kramer, U.M. (2014). Orbitofrontal cortex reactivity to angry facial expression in a social interaction correlates with aggressive behavior. *Cerebral Cortex*, 25(9), 3057–63.
- Blair, R.J. (2007). The amygdala and ventromedial prefrontal cortex in morality and psychopathy. *Trends in Cognitive Sciences*, 11(9), 387–92.
- Blair, R.J. (2010). Psychopathy, frustration, and reactive aggression: the role of ventromedial prefrontal cortex. *British Journal of Psychology*, 101(Pt 3), 383–99.
- Blair, R.J. (2012). Considering anger from a cognitive neuroscience perspective. *Wiley Interdisciplinary Reviews: Cognitive Science*, 3(1), 65–74.
- Buckholtz, J.W., Treadway, M.T., Cowan, R.L., et al. (2010). Mesolimbic dopamine reward system hypersensitivity in individuals with psychopathic traits. *Nature Neuroscience*, 13(4), 419–21.
- Buss, A.H., Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology*, 63(3), 452–9.
- Carre, J.M., Hyde, L.W., Neumann, C.S., Viding, E., Hariri, A.R. (2013). The neural signatures of distinct psychopathic traits. *Society for Neuroscience*, 8(2), 122–35.
- Cherek, D.R., Moeller, F.G., Schnapp, W., Dougherty, D.M. (1997). Studies of violent and nonviolent male parolees: I. Laboratory and psychometric measurements of aggression. *Biological Psychiatry*, 41(5), 514–22.
- Coccaro, E.F., McCloskey, M.S., Fitzgerald, D.A., Phan, K.L. (2007). Amygdala and orbitofrontal reactivity to social threat in individuals with impulsive aggression. *Biological Psychiatry*, 62(2), 168–78.
- Coccaro, E.F., Sripada, C.S., Yanowitch, R.N., Phan, K.L. (2011). Corticolimbic function in impulsive aggressive behavior. *Biological Psychiatry*, 69(12), 1153–9.
- Cooke, D.J., Michie, C., Hart, S.D., Clark, D. (2005). Assessing psychopathy in the UK: concerns about cross-cultural generalisability. *British Journal of Psychiatry*, 186, 335–41.
- da Cunha-Bang, S., Stenbaek, D.S., Holst, K., et al. (2013). Trait aggression and trait impulsivity are not related to frontal cortex 5-HT<sub>2A</sub> receptor binding in healthy individuals. *Psychiatry Research*, 212(2), 125–31.
- da Cunha-Bang, S.H., Perfalk, V.L., Beliveau, E., et al. (2016). Serotonin 1B receptor binding is associated with trait anger and level of psychopathy in violent offenders. *Biological Psychiatry*, doi: 10.1016/j.biopsych.2016.02.030.
- Eklund, A., Nichols, T.E., Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences of the United States of America National Academy of Sciences*, 113(28), 7900–5.
- Forman, S.D., Cohen, J.D., Fitzgerald, M., Eddy, W.F., Mintun, M.A., Noll, D.C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magnetic Resonance in Medicine*, 33(5), 636–47.
- Friston, K.J., Buechel, C., Fink, G.R., Morris, J., Rolls, E., Dolan, R.J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage*, 6(3), 218–29.
- Glenn, A.L., Raine, A., Yaralian, P.S., Yang, Y. (2010). Increased volume of the striatum in psychopathic individuals. *Biological Psychiatry*, 67(1), 52–8.
- Glenn, A.L., Yang, Y. (2012). The potential role of the striatum in antisocial behavior and psychopathy. *Biological Psychiatry*, 72(10), 817–22.
- Gospic, K., Mohlin, E., Fransson, P., Petrovic, P., Johannesson, M., Ingvar, M. (2011). Limbic justice – amygdala involvement in immediate rejection in the Ultimatum Game. *PLoS Biology*, 9(5), e1001054.
- Hare, R. (2003). *Hare Psychopathy Checklist-Revised (PCL-R)*, 2nd edn, Toronto: Multi-Health Systems Inc.
- Madsen, M.K., Mc Mahon, B., Andersen, S.B., Siebner, H.R., Knudsen, G.M., Fisher, P.M. (2015). Threat-related amygdala functional connectivity is associated with 5-HTTLPR genotype and neuroticism. *Social Cognitive and Affective Neuroscience*, 11(1), 140–9.
- McCloskey, M.S., Phan, K.L., Angstadt, M., Fettich, K.C., Keedy, S., Coccaro, E.F. (2016). Amygdala hyperactivation to angry faces in intermittent explosive disorder. *Journal of Psychiatric Research*, 79, 34–41.
- McLaren, D.G., Ries, M.L., Xu, G., Johnson, S.C. (2012). A generalized form of context-dependent psychophysiological interactions (gPPI): a comparison to standard approaches. *Neuroimage*, 61(4), 1277–86.
- Moeller, S.B., Novaco, R.W., Heinola-Nielsen, V., Hougaard, H. (2015). Validation of the Novaco Anger Scale-Provocation Inventory (Danish) with nonclinical, clinical, and offender samples. *Assessment*, 23(5), 624–36.
- Motzkin, J.C., Newman, J.P., Kiehl, K.A., Koenigs, M. (2011). Reduced prefrontal connectivity in psychopathy. *Journal of Neuroscience*, 31(48), 17348–57.
- Nelson, R.J., Trainor, B.C. (2007). Neural mechanisms of aggression. *Nature Reviews Neuroscience*, 8(7), 536–46.
- Patton, J.H., Stanford, M.S., Barratt, E.S. (1995). Factor structure of the Barratt impulsiveness scale. *Journal of Clinical Psychology*, 51(6), 768–74.
- Rosell, D.R., Siever, L.J. (2015). The neurobiology of aggression and violence. *CNS Spectrums*, 20(3), 254–79.
- Schulze, L., Schmah, C., Niedtfeld, I. (2015). Neural correlates of disturbed emotion processing in borderline personality disorder: a multimodal meta-analysis. *Biological Psychiatry*, 79(2), 97–106.
- Siever, L.J. (2008). Neurobiology of aggression and violence. *American Journal of Psychiatry*, 165(4), 429–42.
- Skibsted, A. d., C.-B, S., Carré, J., et al. (2015). Aggression-related brain function assessed with the point subtraction aggression paradigm in functional magnetic resonance imaging. In: *European Neuropsychopharmacology 2015*, Amsterdam.
- White, S.F., Marsh, A.A., Fowler, K.A., et al. (2012). Reduced amygdala response in youths with disruptive behavior disorders and psychopathic traits: decreased emotional response versus increased top-down attention to nonemotional features. *American Journal of Psychiatry*, 169(7), 750–8.
- White, S.F., VanTieghem, M., Brislin, S.J., et al. (2015). Neural correlates of the propensity for retaliatory behavior in youths with disruptive behavior disorders. *American Journal of Psychiatry*, appiajp201515020250.
- Yang, Y., Narr, K.L., Baker, L.A., et al. (2015). Frontal and striatal alterations associated with psychopathic traits in adolescents. *Psychiatry Research*, 231(3), 333–40.
- Yu, R., Mobbs, D., Seymour, B., Rowe, J.B., Calder, A.J. (2014). The neural signature of escalating frustration in humans. *Cortex*, 54, 165–78.
- Zink, C.F., Pagnoni, G., Martin, M.E., Dhamala, M., Berns, G.S. (2003). Human striatal response to salient nonrewarding stimuli. *Journal of Neuroscience*, 23(22), 8092–7.
- Zink, C.F., Pagnoni, G., Martin-Skurski, M.E., Chappel, J.C., Berns, G.S. (2004). Human striatal responses to monetary reward depend on saliency. *Neuron*, 42(3), 509–17.